

Apache Pig - Mise en oeuvre et programmation

Référence : 19003-171017-3-PIG

Durée : 2 jours soit 14 heures

Filière : Décisionnel

Stage disponible en mode « In Class »

Population visée :

Ouvrier – Employé – Employé qualifié – Cadre – Cadre supérieur

Public concerné :

Directeurs techniques, chefs de projets, architectes, consultants, administrateurs de bases de données, développeurs d'applications

Cette formation permet d'acquérir les connaissances et compétences requises pour créer un entrepôt de données Big Data à l'aide de l'outil Apache Hive.

PRÉ-REQUIS

- Avoir des connaissances générales sur le Big Data et l'infrastructure Apache Hadoop.

OBJECTIFS PEDAGOGIQUES

A l'issue de cette formation, l'apprenant sera capable de :

- Alimentation de Hive ;
- Créer des tables ;
- Injecter des données ;
- Réaliser des requêtes, mais aussi optimiser les traitements.

OBJECTIFS OPERATIONNELS ET CONTENU DE LA FORMATION

Présentation du Pig Latin

- Utilisation
- Fonctionnement
- Syntaxe
- Types

Chargement des données

- PigStorage
- Schémas
- Commande DESCRIBE

Sauvegarde de données

- DUMP
- STORE

Autres commandes utiles - cat

Traitement des données

Opérateurs

- FILTER

- ORDER
- GROUP
- FOREACH... GENERATE
- FLATTEN

Fonctions

- Chaînes de caractères
- Mathématiques
- Exemple - Compteur de mots, version Pig

Etendre Pig

- User Defined Functions

Méthodes et moyens :

- Explications théoriques suivies de pratiques guidées puis mises en autonomie
- 1 vidéoprojecteur par salle
- 1 ordinateur par stagiaire

Méthodes d'évaluation des acquis :

- Exercices de synthèse et d'évaluation
- Evaluation de fin de stage

Profil formateur :

Nos formateurs sont certifiés à l'issue d'un parcours organisé par nos soins. Ils bénéficient d'un suivi de compétences aussi bien au niveau technique que pédagogique.

Support stagiaire :

- Support papier ou électronique (dématérialisé)
- Les exercices d'accompagnement peuvent être récupérés sur clef USB